Markov Chains, in general

A vector $v \in \mathbb{R}^n$ is called a **probability vector** if its entries are nonnegative and sum to 1

An $n \times n$ matrix M is a **Markov (or stochastic) matrix** if its columns are probability vectors

Ex: $\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$, $\begin{bmatrix} 1 & .5 & .8 \\ 0 & 0 & .2 \\ 0 & .5 & 0 \end{bmatrix}$

Let $x_0 \in \mathbb{R}^n$ be a probability vector.

Then $x_0, x_1 = Mx_0, x_2 = M^2x_0, \ldots$ is called a **Markov Chain** and x_0 is called the **initial state vector**

Lemma: If x is a probability vector, then so is Mx. Hence every vector in a Markov chain is a probability vector.

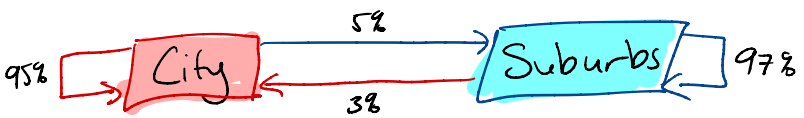
Proof: Let $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ such that $x_1 + \cdots + x_n = 1$

Let $M = [m_1 \cdots m_n]$ be a Markov matrix;
so the entries of m_i sum to 1 $\forall i$.
Notice for any $a \in \mathbb{R}$, the entries of am_i sum to a.
Now the entries of $Mx = x_1m_1 + \cdots + x_nm_n$
sum to $x_1 + \cdots + x_n = 1$

∎

A steady state vector q for M is a probability vector satisfying $Mq = q$ (i.e. an eigenvector with corresponding to eigenvalue $\lambda = 1$)

Ex: Recall the example given by



with $M = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$.

M has eigenvalues $\lambda = 1, .92$ and $E_1(M) = \text{Span}\left\{\begin{bmatrix} .6 \\ 1 \end{bmatrix}\right\}$.

Let q be a steady state vector. then $q = c\begin{bmatrix} .6 \\ 1 \end{bmatrix}$

and $c(.6) + c(1) = 1 \implies c = .625 \implies q = \begin{bmatrix} .375 \\ .625 \end{bmatrix}$ (unique)

Last time we saw that for any initial vector x_0,

$$x_k = M^k x_0 \longrightarrow \begin{bmatrix} .375 \\ .625 \end{bmatrix}$$

Hence every Markov chain converges to the (unique) steady state vector.

We will see that convergence to steady state vectors is a key feature of Markov chains

First we show that steady state vectors always exist

Proposition 1: Let M be a Markov matrix. Then it has a steady state vector.

Proof:

Fact: For any matrix A, A and A^T have the same eigenvalues (recitation)

Now $M^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \text{sum of entries of } 1^{st} \text{ col of } M \\ \vdots \\ \text{sum of entries of } n^{th} \text{ col of } M \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ by def'n

Thus M^T has evector $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ w/ corresp evalue 1.

M has eigenvalue 1. Let $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ be a corresp evector.

then $Mv = v$. It can be shown that $v_i \geq 0 \ \forall i$.

(this is known as the Frobenius-Perron thm)

Hence $q = \frac{1}{v_1 + \cdots + v_n} v$ is a steady state vector ∎

We now prove that $\lambda = 1$ is largest magnitude eigenvalue.

Proposition 2: Let M be a Markov matrix and let λ be an eigenvalue of M. Then $|\lambda| \leq 1$.

Proof:

Assume $|\lambda| > 1$ for some eigenvalue λ of M.

Let $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ be a corresponding unit eigenvector

Hence $\|M^K v\| = \|\lambda^K v\| = |\lambda|^K \|v\| = |\lambda|^K \to \infty$ as $K \to \infty$

Now since $\|v\| = 1$, its entries are $|v_i| \leq 1$

Since entries of M^K are between 0 and 1 and entries of $M^K v$ are dot products of rows of M^K and v, each entry of $M^K v$ is between 0 and n (where M is $n \times n$)

Hence $\|M^K v\| \leq \sqrt{\underbrace{n^2 + \cdots + n^2}_{n}} = \sqrt{n^3} \quad \forall K$.

But this contradicts $\|M^K v\| \to \infty$

Hence $|\lambda| \leq 1$. ∎

Theorem: Let M be a diagonalizable Markov matrix with a unique eigenvalue λ such that $|\lambda| = 1$. Then there is a unique steady state vector q and every Markov chain $x_K = M^K x_0$ converges to q

Proof:

By Prop 1, $\lambda = 1$ is an eigenvalue of M.

By hypothesis, $\lambda = 1$ has algebraic multiplicity 1.

Hence $\dim E_1(M) = 1$.

Suppose q_1 and q_2 are steady state vectors.

Then $q_1 = cq_2$ for some $c \in \mathbb{R}$.

Since the entries of q_1 add to 1, so do the entries of cq_2.

But, the entries of q_2 add to 1 and so the entries of cq_2 add to c. Hence $c = 1$ and so $q_1 = q_2$

Thus, there is a <u>unique</u> steady state vector q.

Since M is diagonalizable, it has n linearly independent eigenvectors $v_1, \dots, v_n$. Let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues.

By Prop. 1, we may assume $\lambda_1 = 1$ and $v_1 = q$.

By hypothesis, $|\lambda_i| < 1$ for all $2 \le i \le n$.

Since $\{q, v_2, \dots, v_n\}$ is linearly independent, it is a basis of $\mathbb{R}^n$.

Thus $x_0 = c_1 q + c_2 v_2 + \cdots + c_n v_n$ for some $c_1, \dots, c_n \in \mathbb{R}$.

Now
$$\begin{aligned} x_k = M^k x_0 &= M^k (c_1 q + c_2 v_2 + \cdots + c_n v_n) \\ &= c_1 M^k q + c_2 M^k v_2 + \cdots + c_n M^k v_n \\ &= c_1 q + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n \end{aligned}$$

As $k \to \infty$, $\lambda_2^k, \dots, \lambda_n^k \to 0$ since $|\lambda_i| < 1 \ \forall \ 2 \le i \le n$

Hence as $k \to \infty$, $x_k \to c_1 q$

Let $x_k = \begin{bmatrix} x_1^k \\ \vdots \\ x_n^k \end{bmatrix}$ and $c_1 q = \begin{bmatrix} c_1 q_1 \\ \vdots \\ c_1 q_n \end{bmatrix}$.

Then $x_i^k \to c_1 q_i \ \forall \ i$ as $k \to \infty$

Hence $x_1^k + \cdots + x_n^k \to c_1 q_1 + \cdots + c_1 q_n = c_1$ as $k \to \infty$

Since $x_1^k + \cdots + x_n^k = 1 \ \forall \ k$ since $M^k x_0$ is a prob. vector,

It follows that $c_1 = 1$.

Hence $x_k \to q$ as $k \to \infty$

A Markov matrix is called positive if all entries of M are positive

E.g. $\begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$ is positive

Fact: Every positive Markov matrix satisfies the hypotheses of the previous theorem and so:

Theorem: If M is a positive Markov matrix, then it has a unique steady state vector q and every Markov Chain converges to q.

More generally, a Markov matrix M is called regular if M^K is positive for some $K>0$.

Such matrices can have complex eigenvalues. However, the theorem still holds!

Theorem: If M is a regular Markov matrix, then it has a unique steady state vector q and every Markov chain converges to q.

Intuition: For some K, M^K is positive.
Hence by the previous theorem, $M^K x_0$ converges to a unique steady state $\forall\, x_0$.
So the same should be true for M.
The full proof uses the strategy we used above but with complex evalues and evectors.